

Pinching Antenna System–Assisted Hybrid AirComp–NOMA Uplink: Joint Precoding and Antenna Placement Optimization

Saeid Pakravan, Imene Trigui, Wessam Ajib, and Wei-Ping Zhu

Abstract—This paper studies a pinching antenna system (PAS)-assisted hybrid uplink architecture that integrates over-the-air computation (AirComp) and non-orthogonal multiple access (NOMA) to simultaneously support distributed data aggregation and individual communication services. A base station with a dielectric waveguide hosting multiple pinching antennas receives signals from AirComp and NOMA users over shared time–frequency resources. To assess joint computation–communication performance, a hybrid metric combining the AirComp computation rate and the NOMA sum rate is proposed. Based on this metric, a joint optimization problem is formulated to maximize the hybrid rate by optimizing user transmit precoding, receive combining, and antenna deployment, subject to power, quality-of-service, and aggregation accuracy constraints. An alternating optimization framework is developed to solve the resulting non-convex problem. Numerical results show that the proposed design achieves significant performance gains over several benchmark schemes.

Index Terms—Pinching antenna systems, over-the-air computation, non-orthogonal multiple access, alternating optimization.

I. INTRODUCTION

The rapid evolution toward sixth-generation (6G) wireless networks, coupled with the explosive growth of the Internet of Things (IoT), is driving unprecedented demands for ultra-low latency, massive connectivity, and high spectral efficiency [1]. Emerging 6G applications, such as immersive extended reality, autonomous systems, and large-scale distributed intelligence, demand not only reliable data delivery but also native support for computation within the wireless infrastructure [2]. Conventional transmit-then-compute paradigms, where data are first decoded and then processed, face fundamental scalability and latency limitations in dense multiple-access scenarios.

Over-the-air computation (AirComp) has recently emerged as a promising solution. By exploiting the natural superposition property of the wireless multiple-access channel (MAC), AirComp enables the computation of functions of distributed data directly in the air [3]. This simultaneous transmission and aggregation approach reduces communication latency and signaling overhead, making it attractive for federated learning, distributed sensing, and edge intelligence [4]–[6]. In

parallel, non-orthogonal multiple access (NOMA) has been widely studied to enhance spectral efficiency and support massive connectivity, allowing multiple users to share the same time–frequency resources via power-domain multiplexing and successive interference cancellation (SIC) [7].

Despite their advantages, AirComp and NOMA have largely been studied in isolation, as they are traditionally designed to serve different objectives. Specifically, AirComp focuses on the accuracy of function computation, whereas NOMA aims to maximize individual data rates. In future 6G networks, however, heterogeneous service coexistence is unavoidable, where computation-oriented devices (e.g., sensing and learning nodes) and communication-oriented users (e.g., mobile broadband terminals) operate simultaneously over shared wireless resources [8]. This coexistence introduces challenges in interference management, resource allocation, and the tradeoff between computation fidelity and communication throughput.

Recent advances in reconfigurable antenna technologies offer additional spatial degrees of freedom (DoF) to address these challenges. Among these technologies, pinching antenna systems (PAS) have emerged as a novel architecture in which the effective antenna position can be continuously adjusted along a conductive or dielectric medium via controlled pinching mechanisms [9]–[12]. Unlike fixed-position antennas, PAS can dynamically adapt to channel conditions and user distributions. This flexibility enables joint spatial and signal-domain optimization, which is especially beneficial in interference-limited integrated communication–computation systems.

This paper studies a PAS-enabled uplink framework supporting two types of users: AirComp users, which transmit function-related signals for over-the-air aggregation, and NOMA users, which seek reliable and high-rate data transmission. A base station (BS) equipped with PAS simultaneously aggregates AirComp signals while decoding individual data streams from NOMA users over shared spectrum resources. The spatial flexibility offered by PAS is leveraged to reshape the effective channels, thereby balancing the conflicting requirements of computation accuracy and communication rate.

The main objective of this work is to optimize overall system performance using a hybrid rate metric that jointly captures NOMA data rates and AirComp computation efficiency. The resulting optimization problem is highly non-convex, involving coupled variables such as antenna pinching positions, transmit power allocation, and receiver design, under practical constraints like power budgets and decoding requirements. To solve this, we develop an alternating optimization (AO) framework that decomposes the original prob-

This work was supported by the Regroupement Stratégique en Microsystèmes du Québec (ReSMiQ).

S. Pakravan and W. Ajib are with the Department of Computer Sciences, University of Quebec in Montreal (UQAM), Montreal, QC, Canada. email: pakravan.saeid@uqam.ca; ajib.wessam@uqam.ca.

I. Trigui is with the Department of Applied Sciences, University of Quebec at Chicoutimi (UQAC), Chicoutimi, QC, Canada. email: itrigui@uqac.ca.

W.-P. Zhu is with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. email: weiping@ece.concordia.ca.

lem into tractable subproblems, solved iteratively. Numerical results show that the proposed design significantly outperforms benchmark schemes, achieving higher hybrid rates while providing valuable insights into the tradeoff between computation accuracy and communication throughput.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

We consider a PAS-aided uplink multiple-access system, where the BS is equipped with a dielectric waveguide hosting N PAs. The system simultaneously serves two disjoint classes of single-antenna users: a set of AirComp users denoted by $\mathcal{K}_A \triangleq \{1, \dots, K_A\}$ and a set of uplink NOMA users denoted by $\mathcal{K}_N \triangleq \{1, \dots, K_N\}$, with $\mathcal{K}_A \cap \mathcal{K}_N = \emptyset$. The overall user set is given by $\mathcal{U} \triangleq \mathcal{K}_A \cup \mathcal{K}_N$. The BS aggregates wireless data from the AirComp users while decoding individual data streams from NOMA users over the same time-frequency resources. We adopt a three-dimensional (3D) Cartesian co-ordinate system. The users are uniformly distributed over a rectangular area on the x - y plane with side lengths L_x and L_y . The position of user $u \in \mathcal{U}$ is denoted by $\mathbf{u}_u = (x_u, y_u, 0)$. Without loss of generality, the dielectric waveguide is deployed along the x -axis at a height d above the ground, with length L_x . The position of the n -th PA on the waveguide is given by $\mathbf{v}_n = (v_n, 0, d)$, $0 \leq v_n \leq L_x$, while the feed point of the waveguide is located at $\mathbf{v}_0 = (0, 0, d)$.

The channel coefficient between user u and the n -th PA is modeled as [13]

$$h_{u,n} = \frac{\lambda}{4\pi\|\mathbf{u}_u - \mathbf{v}_n\|} \exp\left(-j\frac{2\pi}{\lambda}\|\mathbf{u}_u - \mathbf{v}_n\|\right), \quad (1)$$

where λ denotes the carrier wavelength.

Let $s_k \in \mathbb{C}$ and $d_j \in \mathbb{C}$ represent the data symbols transmitted by AirComp user $k \in \mathcal{K}_A$ and NOMA user $j \in \mathcal{K}_N$, respectively. Accordingly, the received signal at the n -th PA is expressed as

$$r_n = \sum_{k \in \mathcal{K}_A} h_{k,n} p_k s_k + \sum_{j \in \mathcal{K}_N} h_{j,n} p_j d_j + z_n, \quad (2)$$

where p_k and p_j are the transmit precoding coefficients of AirComp and NOMA users, respectively, and $z_n \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian noise (AWGN).

For analytical tractability, all symbols are assumed to be zero-mean with unit power, i.e., $\mathbb{E}[s_k] = 0$, $\mathbb{E}[|s_k|^2] = 1$, $\forall k \in \mathcal{K}_A$, $\mathbb{E}[d_j] = 0$, $\mathbb{E}[|d_j|^2] = 1$, $\forall j \in \mathcal{K}_N$. Furthermore, the transmitted symbols are mutually independent, $\mathbb{E}[s_k s_i^H] = 0$, $\forall k \neq i$, $\mathbb{E}[d_j d_i^H] = 0$, $\forall j \neq i$. Then, the per-user transmit power constraint is expressed as

$$\mathbb{E}[|p_u x_u|^2] = |p_u|^2 \leq P_u, \quad \forall u \in \mathcal{U}, \quad (3)$$

where $x_u \in \mathbb{C}$ denotes the information symbol of user u , i.e., $x_u = s_k$ for $u \in \mathcal{K}_A$ and $x_u = d_j$ for $u \in \mathcal{K}_N$, and $P_u > 0$ represents the maximum transmit power budget of user u .

Assuming a fully passive waveguide structure, the aggregated signal observed at the feed point is given by

$$\tilde{r} = \sum_{n=1}^N r_n \exp\left(-j\frac{2\pi n_{\text{ref}}}{\lambda}\|\mathbf{v}_n - \mathbf{v}_0\|\right), \quad (4)$$

where n_{ref} denotes the refractive index of the waveguide. Defining the equivalent uplink channel as

$$g_u \triangleq \sum_{n=1}^N h_{u,n} \exp\left(-j\frac{2\pi n_{\text{ref}}}{\lambda}\|\mathbf{v}_n - \mathbf{v}_0\|\right), \quad (5)$$

the received signal can be compactly written as

$$\tilde{r} = \sum_{k \in \mathcal{K}_A} g_k p_k s_k + \sum_{j \in \mathcal{K}_N} g_j p_j d_j + \tilde{z}, \quad (6)$$

where $\tilde{z} \sim \mathcal{CN}(0, N\sigma^2)$.

B. NOMA Decoding via SIC and Achievable Rates

Since all users are served over the same bandwidth simultaneously, interference management is critical in the considered hybrid AirComp-NOMA network. To mitigate intra-NOMA interference and enable subsequent AirComp aggregation, the BS employs SIC at the feed point of the PAS. The NOMA users are indexed in descending order of their equivalent channel gains as $|g_1|^2 \geq |g_2|^2 \geq \dots \geq |g_{K_N}|^2$, where user 1 has the strongest channel and is decoded first, while user K_N has the weakest channel and is decoded last. At the decoding stage of NOMA user j , the signals of stronger NOMA users $\{1, \dots, j-1\}$ have already been decoded and perfectly canceled via SIC, whereas the signals of weaker NOMA users $\{j+1, \dots, K_N\}$ remain as intra-NOMA interference. Moreover, AirComp users transmit functional symbols intended for over-the-air aggregation rather than individually decodable messages. Hence, their signals cannot be decoded and canceled during the NOMA SIC process and are treated as non-removable interference at all NOMA decoding stages, whereas the reconstructed waveforms of successfully decoded NOMA users can be removed via SIC.

Based on this SIC structure, the SINR for decoding NOMA user j is given by

$$\Gamma_j = \frac{|g_j|^2 |p_j|^2}{\sum_{t=j+1}^{K_N} |g_t|^2 |p_t|^2 + \sum_{k \in \mathcal{K}_A} |g_k|^2 |p_k|^2 + N\sigma^2}. \quad (7)$$

Accordingly, the achievable uplink rate of NOMA user j is

$$R_{N,j} = B \log_2(1 + \Gamma_j), \quad (8)$$

and the NOMA sum rate is $R_N = \sum_{j=1}^{K_N} R_{N,j}$.

C. AirComp Estimation and Computation Rate

After all NOMA users have been successfully decoded according to the above SIC order, the BS reconstructs and subtracts their signals from the received superposed signal. The remaining signal contains only the AirComp components and noise, given by

$$\hat{r} = \sum_{k \in \mathcal{K}_A} g_k p_k s_k + \tilde{z}. \quad (9)$$

The goal of AirComp is to estimate $s \triangleq \sum_{k \in \mathcal{K}_A} s_k$. The BS applies a scalar receive combining coefficient $w \in \mathbb{C}$ to obtain

$$\hat{s} = w^H \sum_{k \in \mathcal{K}_A} g_k p_k s_k + w^H \tilde{z}. \quad (10)$$

The computation distortion of $\hat{s}$ with respect to s is quantified by the MSE, defined as $\text{MSE}(\hat{s}, s) \triangleq \mathbb{E} \left[|\hat{s} - s|^2 \right]$, given by

$$\text{MSE}(\hat{s}, s) = \sum_{k \in \mathcal{K}_A} |w^H g_k p_k - 1|^2 + N\sigma^2 |w|^2. \quad (11)$$

Since it directly measures the computation distortion, it serves as a key performance metric for collaborative computation. Accordingly, the achievable computation rate is defined as [14]

$$R_A \triangleq B \log_2 \left(1 + \frac{\mathbb{E} [|\hat{s}|^2] - \text{MSE}(\hat{s}, s)}{\text{MSE}(\hat{s}, s)} \right). \quad (12)$$

where $\mathbb{E} [|\hat{s}|^2] = \sum_{k \in \mathcal{K}_A} |w^H g_k p_k|^2 + N\sigma^2 |w|^2$.

D. Hybrid Performance Metric

To jointly capture AirComp accuracy and NOMA throughput, we define the hybrid rate as

$$R_H \triangleq \alpha R_A + (1 - \alpha) R_N, \quad (13)$$

where $\alpha \in [0, 1]$ is a weighting parameter that controls the relative importance of AirComp aggregation fidelity and NOMA communication throughput. The proposed hybrid metric enables unified evaluation of heterogeneous uplink services and reduces to AirComp optimization or conventional NOMA sum-rate maximization when $\alpha = 1$ or $\alpha = 0$, respectively.

E. Problem Formulation

We aim to maximize the achievable hybrid rate in (13) by jointly optimizing the users' transmit precoding coefficients, the deployment of the PAs along the dielectric waveguide, and the receive combining coefficient at the BS. Accordingly, the hybrid performance maximization problem is formulated as

$$\max_{w, \mathbf{p}_k, \mathbf{p}_j, \mathbf{v}} R_H(w, \mathbf{p}_k, \mathbf{p}_j, \mathbf{v}) \quad (14a)$$

$$\text{s.t. } |p_k|^2 \leq P_k, \quad \forall k \in \mathcal{K}_A, \quad (14b)$$

$$|p_j|^2 \leq P_j, \quad \forall j \in \mathcal{K}_N, \quad (14c)$$

$$R_{N,j} \geq R_{\min,j}, \quad \forall j \in \mathcal{K}_N, \quad (14d)$$

$$\text{MSE}(\hat{s}, s) \leq \varepsilon_0, \quad (14e)$$

$$0 \leq v_n \leq L_x, \quad \forall n = 1, \dots, N, \quad (14f)$$

where $\mathbf{p}_k \triangleq [p_k]_{k \in \mathcal{K}_A}$, $\mathbf{p}_j \triangleq [p_j]_{j \in \mathcal{K}_N}$, and $\mathbf{v} \triangleq [v_1, \dots, v_N]^T$ represents the PA locations along the dielectric waveguide. Constraints (14b) and (14c) impose individual transmit power budgets of the AirComp and NOMA users, respectively. Constraint (14d) guarantees the quality-of-service (QoS) requirements of the NOMA users, where $R_{\min,j}$ denotes the minimum achievable rate of user j . The AirComp accuracy requirement is captured by (14e), which constrains the aggregation MSE to be no larger than a predefined threshold. Finally, constraint (14f) ensures that all PAs are deployed within the physical extent of the dielectric waveguide.

Problem (14) is non-convex due to coupled variables and nonlinear dependence of both the hybrid objective function and the QoS constraints on the PAS-aided equivalent channels. This motivates the development of efficient joint optimization algorithms for precoding design, PA placement, and receive combining, which will be addressed in the subsequent sections.

III. PROPOSED SOLUTION

We employ an AO framework to decompose the problem into tractable subproblems, where transmit powers, receive scalar, and PA location vector are optimized alternately with other variables fixed.

A. Transmit Power Allocation

For given w and $\mathbf{v}$, define the equivalent PAS-assisted channels as $\bar{g}_k \triangleq g_k(\mathbf{v})$. Due to SIC decoding order, we first optimize NOMA transmit powers and then AirComp powers.

1) *Power Allocation of NOMA Users*: For fixed $\mathbf{p}_k$, the NOMA sum-rate can be rewritten using the telescoping property of SIC as

$$R_N = B \log_2 \left(1 + \frac{\sum_{j \in \mathcal{K}_N} |p_j|^2 |\bar{g}_j|^2}{\sum_{k \in \mathcal{K}_A} |p_k|^2 |\bar{g}_k|^2 + N\sigma^2} \right). \quad (15)$$

Let $\rho_j = |p_j|^2$ and define $I_A = \sum_{k \in \mathcal{K}_A} |p_k|^2 |\bar{g}_k|^2$. The QoS constraint $R_j^N \geq R_j^{\min}$ can be equivalently rewritten as

$$\rho_j |\bar{g}_j|^2 \geq \gamma_j \left(\sum_{t=j+1}^{K_N} \rho_t |\bar{g}_t|^2 + I_A + N\sigma^2 \right), \quad (16)$$

where $\gamma_j = 2^{R_{\min,j}/B} - 1$. Therefore, the NOMA power allocation subproblem is given by

$$\max_{\{\rho_j\}} (1 - \alpha) B \log_2 \left(1 + \frac{\sum_j \rho_j |\bar{g}_j|^2}{I_A + N\sigma^2} \right) \quad (17a)$$

$$\text{s.t. } \rho_j |\bar{g}_j|^2 \geq \gamma_j \left(\sum_{t=j+1}^{K_N} \rho_t |\bar{g}_t|^2 + I_A + N\sigma^2 \right), \quad (17b)$$

$$0 \leq \rho_j \leq P_j, \quad \forall j \in \mathcal{K}_N. \quad (17c)$$

The objective function is concave and all constraints are linear in $\{\rho_j\}$, hence this problem is convex and can be efficiently solved by the standard convex optimization tools such as CVX [15]. The optimal transmit powers are recovered as $p_j^* = \sqrt{\rho_j^*}$.

2) *Power Allocation of AirComp Users*: With $\mathbf{p}_j^*$ obtained, the AirComp power allocation problem is considered. Although the feasible set is convex, the objective function is non-concave. To address this issue, we exploit its difference-of-convex (DC) structure and solve it within the DC programming framework [16]. Define

$$F(\mathbf{p}_k) = (1 - \alpha) B \log_2 \left(\sum_{k \in \mathcal{K}_A} |p_k|^2 |\bar{g}_k|^2 + I_N + N\sigma^2 \right) + \alpha B \log_2 \left(\sum_{k \in \mathcal{K}_A} |w^H \bar{g}_k p_k|^2 + N\sigma^2 |w|^2 \right), \quad (18)$$

$$G(\mathbf{p}_k, \beta) = (1 - \alpha) B \log_2 \left(\sum_{k \in \mathcal{K}_A} |p_k|^2 |\bar{g}_k|^2 + N\sigma^2 \right) + \alpha B \log_2 (\beta + N\sigma^2 |w|^2), \quad (19)$$

where $I_N = \sum_{j \in \mathcal{K}_N} |p_j^*|^2 |\bar{g}_j|^2$ is constant. An auxiliary variable β is introduced to satisfy $\beta \geq \sum_{k \in \mathcal{K}_A} |w^H \bar{g}_k p_k - 1|^2$, which preserves convexity of the feasible set. Then, the hybrid objective can be rewritten as $R_H(\mathbf{p}_k) = F(\mathbf{p}_k) - G(\mathbf{p}_k, \beta)$, which admits a DC decomposition. At iteration ℓ , the convex

function $G(\mathbf{p}_k, \beta)$ is approximated by its first-order Taylor expansion at $(\mathbf{p}_k^{(\ell)}, \beta^{(\ell)})$, resulting in a convex problem that can be efficiently solved. By iteratively updating the solution, the DC algorithm generates a sequence of feasible points with non-decreasing objective values and converges to a stationary point.

B. Receive Scalar Optimization

For fixed transmit powers and PA locations, the receive scalar optimization is given by

$$\max_w \log_2 \left(\frac{\sum_k |w^H \bar{g}_k p_k|^2 + N\sigma^2 |w|^2}{\sum_k |w^H \bar{g}_k p_k - 1|^2 + N\sigma^2 |w|^2} \right) \quad (20)$$

subject to the MSE constraint.

Introduce $\tilde{w} = 1/w$. The problem is equivalent to

$$\min_{\tilde{w}} \sum_{k \in \mathcal{K}_A} |\tilde{w} - \bar{g}_k p_k|^2 \quad (21a)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}_A} |\tilde{w} - \bar{g}_k p_k|^2 + N\sigma^2 \leq \varepsilon_0 |\tilde{w}|^2. \quad (21b)$$

The objective is convex in $\tilde{w}$, whereas the constraint is non-convex due to the quadratic term $|\tilde{w}|^2$ on the right-hand side. To address this issue, we adopt successive convex approximation (SCA) by linearizing $|\tilde{w}|^2$ via its first-order Taylor expansion at the current iterate. The resulting problem is convex and can be efficiently solved at each iteration, yielding a sequence that converges to a stationary point. The optimal receive scalar is recovered as $w^* = 1/\tilde{w}^*$.

C. Optimization of PA Locations

In the proposed system, the equivalent channels $g_k(\mathbf{v})$ vary continuously with the PA locations $\mathbf{v}$; however, the non-convex dependence induced by pathloss and phase accumulation renders problem (14) intractable with respect to $\mathbf{v}$, motivating the adoption of a projected gradient ascent (PGA) method [17].

1) *Projected Gradient Solution:* The hybrid rate R_H depends on $\mathbf{v}$ through the equivalent channels

$$g_u(\mathbf{v}) = \sum_{n=1}^N h_{u,n}(v_n) e^{-j \frac{2\pi n_{\text{ref}}}{\lambda} v_n}. \quad (22)$$

Let $r_{u,n} = \|\mathbf{u}_u - \mathbf{v}_n\|$. The derivative of g_k with respect to v_n is obtained via the chain rule as

$$\frac{\partial g_u}{\partial v_n} = \left(\frac{\partial h_{u,n}}{\partial v_n} - j \frac{2\pi n_{\text{ref}}}{\lambda} h_{u,n} \right) e^{-j \frac{2\pi n_{\text{ref}}}{\lambda} v_n}, \quad (23)$$

where $\frac{\partial h_{u,n}}{\partial v_n} = h_{u,n} \left(-\frac{v_n - x_u}{r_{u,n}^2} - j \frac{2\pi}{\lambda} \frac{v_n - x_u}{r_{u,n}} \right)$. Using the above expression, the gradient of R_H with respect to v_n is computed as

$$\frac{\partial R_H}{\partial v_n} = \alpha \frac{\partial R_A}{\partial v_n} + (1 - \alpha) \sum_{j=1}^{K_N} \frac{\partial R_j^N}{\partial v_n}, \quad (24)$$

where each term is obtained via standard differentiation of logarithmic rate functions with respect to $|g_k|^2$. Then, we update the PA locations via PGA method:

$$v_n^{(t+1)} = \Pi_{[0, L_x]} \left(v_n^{(t)} + \eta^{(t)} \frac{\partial R_H}{\partial v_n} \right), \quad (25)$$

where $\eta^{(t)}$ is determined using backtracking line search and $\Pi_{[0, L_x]}(\cdot)$ denotes projection onto $[0, L_x]$. After each PA-location update, the equivalent channels are recomputed and the NOMA users are re-indexed according to the adopted descending channel-gain SIC order.

Since $R_H(\mathbf{v})$ is continuously differentiable and the feasible set is compact and convex, the PGA method with appropriate step sizes generates a sequence whose limit points satisfy the first-order KKT conditions. Hence, the algorithm converges to a stationary solution.

D. Convergence and Complexity Analysis

1) *Convergence:* Let $R_H^{(t)}$ denote the objective value at the t -th iteration. Since each block update does not decrease the objective, the sequence $\{R_H^{(t)}\}$ is monotonically non-decreasing. Moreover, the objective is upper bounded owing to the finite transmit power constraints and the bounded PA deployment region. Therefore, $\{R_H^{(t)}\}$ converges. By standard block coordinate descent arguments, the proposed AO algorithm converges to a stationary point of problem (14).

2) *Complexity:* The convex NOMA power allocation with K_N variables requires $\mathcal{O}(K_N^3)$ complexity via interior-point methods. The AirComp power allocation is solved using DC programming, leading to $\mathcal{O}(T_{\text{DC}} K_A^3)$ complexity, where T_{DC} is the number of DC iterations. The receive scalar update incurs negligible cost, while the PA location optimization requires $\mathcal{O}(T_v (K_A + K_N) N)$ operations with T_v projected gradient iterations. Overall, with T_{AO} outer iterations, the total complexity is $\mathcal{O}(T_{\text{AO}} (K_N^3 + T_{\text{DC}} K_A^3 + T_v (K_A + K_N) N))$.

IV. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed PAS-assisted hybrid NOMA–AirComp system. The system parameters are set as follows [12], [16]–[18]. The carrier frequency is set as $f_c = 28$ GHz. The system bandwidth is $B = 1$ MHz. The noise power is assumed to be $\sigma^2 = -90$ dBm. The dielectric waveguide length is $L_x = 10$ m and deployed at height $d = 5$ m. Users are uniformly distributed over a rectangular area of size 10×10 m². Unless otherwise stated, we set $K_A = 4$, $K_N = 3$, $N = 6$, $\alpha = 0.5$, and identical maximum transmit powers $P_k = P_j = P_{\text{max}} = 20$ dBm. The minimum rate requirement of each NOMA user is $R_{\text{min},j} = 0.5$ Mbps and the AirComp MSE threshold is $\varepsilon_0 = 10^{-2}$. All results are averaged over 200 independent user location realizations. We compare the proposed design with the following benchmark schemes:

- **Fixed PA position:** PAs are uniformly distributed on the waveguide with fixed locations given by $v_n = \frac{nL_x}{N+1}$. Only transmit powers and receive scalar are optimized.
- **Discrete PAS:** Each antenna can only be selected from a predefined discrete grid of positions along the waveguide.
- **Full power transmission:** All users transmit at their maximum power P_{max} , while PA locations and receive scalar are optimized.

Fig. 1a illustrates the convergence behavior of hybrid rate versus the number of AO iterations. It can be observed that the hybrid rate increases rapidly with the number of iterations

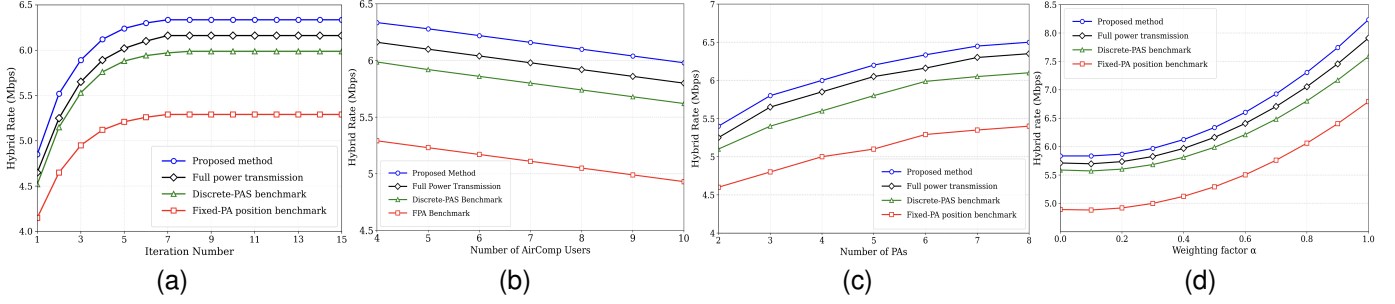


Fig. 1. (a) Convergence performance of our proposed design. (b) Hybrid rate w.r.t. the number of AirComp users. (c) Hybrid rate w.r.t. the number of PAs. (d) Hybrid rate versus the weighting factor α .

and converges within approximately 8 iterations. Moreover, the proposed method consistently achieves a higher hybrid rate compared with the benchmark schemes, demonstrating the effectiveness of the proposed joint optimization framework.

Fig. 1b shows the hybrid rate as a function of the number of AirComp users K_A , while fixing $K_N = 3$ and $N = 6$. As K_A increases, more computation signals share the channel and act as non-removable interference during all NOMA decoding stages, while also increasing the difficulty of accurate AirComp aggregation. Therefore, the hybrid rate gradually decreases for all schemes. Nevertheless, the proposed PAS design consistently achieves a higher hybrid rate than the benchmark schemes, due to the additional spatial DoF that enable more effective shaping of the equivalent channels.

Fig. 1c presents the hybrid rate versus the number of pinching antennas N , with $K_A = 4$ and $K_N = 3$. It is observed that increasing N improves the hybrid rate for all schemes due to enhanced spatial diversity and array gain. However, the performance gap between the proposed design and the benchmark schemes enlarges as N increases. This is because a larger number of antennas provides more spatial DoF, which the proposed continuous optimization fully exploits, whereas the fixed-position scheme cannot adapt its geometry and D-PAS is constrained by discrete positioning resolution.

Fig. 1d illustrates the impact of the weighting factor α on the hybrid rate, which governs the tradeoff between AirComp computation rate and NOMA sum-rate. As α increases, the system progressively shifts from a communication-oriented regime toward a computation-oriented regime. In the low- α region, the hybrid performance is dominated by the NOMA component, resulting in relatively similar performance across all schemes. However, as α increases, the proposed PAS-assisted design achieves noticeably higher hybrid rates compared to the benchmarks. This gain stems from its ability to jointly optimize antenna placement and precoding, thereby effectively shaping the equivalent channels to reduce aggregation distortion while controlling interference.

V. CONCLUSION

This paper investigated a PAS-assisted hybrid framework integrating AirComp and NOMA. A hybrid rate maximization problem was formulated by jointly optimizing transmit power, the AirComp receive scalar, and the positions of pinching antennas. An AO algorithm was developed to solve the resulting non-convex problem. Simulation results demonstrate

that the proposed design achieves significant performance improvements over benchmark schemes.

REFERENCES

- [1] C.-X. Wang *et al.*, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 905–974, Feb. 2023.
- [2] M. Z. Chowdhury *et al.*, "6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 957–975, Jul. 2020.
- [3] Z. Wang *et al.*, "Over-the-air computation for 6G: Foundations, technologies, and applications," *IEEE Internet Things J.*, vol. 11, no. 14, pp. 24 634–24 658, Jul. 2024.
- [4] M. Ahmadzadeh *et al.*, "AI-based fluid antenna design for client selection in over-the-air federated learning," *IEEE Internet Things J.*, vol. 12, no. 20, pp. 42 549–42 558, Oct. 2025.
- [5] G. Zhu *et al.*, "Over-the-air computing for wireless data aggregation in massive IoT," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 57–65, Sep. 2021.
- [6] S. Pakravan *et al.*, "Robust resource allocation for over-the-air computation networks with fluid antenna array," in *Proc. IEEE Globecom Workshops*, Cape Town, South Africa, pp. 1–6, Aug. 2024.
- [7] M. Zeng *et al.*, "Energy-efficient power allocation in uplink mmWave massive MIMO with NOMA," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3000–3004, Mar. 2019.
- [8] W. Saad *et al.*, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, Jun. 2020.
- [9] Z. Ding *et al.*, "Flexible-antenna systems: A pinching-antenna perspective," *IEEE Trans. Commun.*, vol. 73, no. 10, pp. 9236–9253, Oct. 2025.
- [10] S. Pakravan *et al.*, "Impact of position uncertainty on the secrecy performance of pinching-antenna systems," *IEEE Commun. Lett.*, vol. 30, pp. 2765–2769, Jul. 2026.
- [11] X. Xu *et al.*, "Pinching-antenna systems (PASS): Power radiation model and optimal beamforming design," *IEEE Trans. Commun.*, vol. 74, pp. 2160–2175, Nov. 2025.
- [12] S. Pakravan *et al.*, "AI-empowered resource allocation for wirelessly powered pinching-antenna systems," *IEEE Trans. Veh. Technol.*, pp. 1–6, Apr. 2026.
- [13] Y. Li *et al.*, "Pinching antenna-aided wireless powered communication networks," *IEEE Wireless Commun. Lett.*, vol. 15, pp. 255–259, Oct. 2025.
- [14] Y. Hu *et al.*, "Energy minimization for federated learning with IRS-assisted over-the-air computation," in *Proc. IEEE ICASSP*, Toronto, Canada, pp. 3105–3109, Jun. 2021.
- [15] M. Grant *et al.*, "CVX: Matlab software for disciplined convex programming, version 2.1," [Online]. Available: <http://cvxr.com/cvx>, Mar. 2014.
- [16] K. Yang *et al.*, "Federated learning via over-the-air computation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2022–2035, Jan. 2020.
- [17] M. A. Dastgheib *et al.*, "Optimal placement of access points in cellular visible light communication networks: An adaptive gradient projection method," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6813–6825, Jul. 2020.
- [18] Z. Lyu *et al.*, "Pinching-antenna systems (PASS) aided over-the-air computation," *IEEE Commun. Lett.*, vol. 29, no. 11, pp. 2531–2535, Aug. 2025.